NWP test suite and COSMO-LEPS upgrade

Andrea Montani,
Arpae Emilia-Romagna Servizio IdroMeteoClima, Bologna, Italy

COSMO WG6

Outline

- Introduction to COSMO NWP test suite
- Present status and open issues
- Status of migration to 20-member COSMO-LEPS:
  - meteorological aspects
  - computational aspects

NWP test suite

Goals of NWP test suite

- Build up a software environment to perform carefully-controlled and rigorous testing with calculation of verification statistics for any COSMO model test - version
- Offer necessary information on the model forecasting performance
- Provide the COSMO community with standards against which the impacts of new developments in the model should be evaluated.
- Benchmark to monitor the progress of mesoscale forecast improvement (periodic testing as COSMO evolves).

Old status of the suite (installed at ECMWF)

- The suite is implemented to test the present version of COSMO (e.g v5.01) and the experimental one (e.g. v5.03) for 2 months (January 2013 and July 2013) **at 7 km (40 ML, fc+72h, starting at 00UTC) .**
- Both initial and boundary conditions are provided by ECMWF HRES (no nudging):

HRES ➔ COSMO@7p0

- As for observations, synop reports from a domain covering most of Europe and the Middle East are used (**about 3600 stations x day**).
- Output fields are stored on ECMWF ecfs and provided to Versus (also installed at ECMWF) for the comparison of the 2 model versions with the computation of scores and plots.

New status of the suite (installed at ECMWF)

- The suite has been upgraded to test the present version of COSMO (e.g v5.03) and the experimental one (e.g. v5.04a) for 2 months (January 2013 and July 2013) **at both 7 km (40 ML, fc+72h) and 2.8 km (50 ML, fc+48h), always starting at 00UTC.**
- Both initial and boundary conditions are provided by ECMWF HRES (no nudging):

HRES ➔ COSMO@7p0

HRES ➔ COSMO@2p8

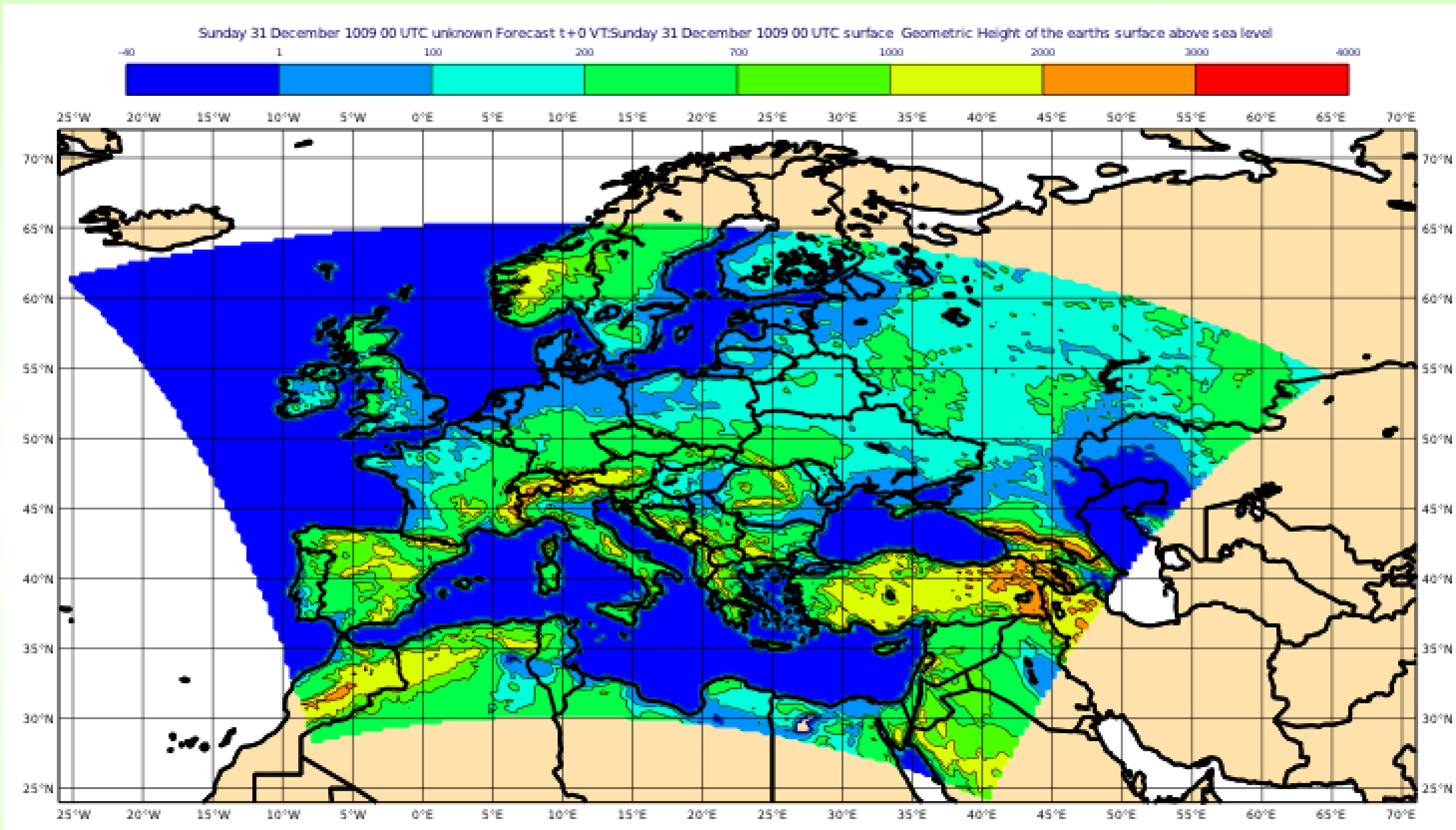
- As for observations, synop reports from a domain covering most of Europe and the Middle East are used (**about 3600 stations x day**).
- Output fields are stored on ECMWF ecfs and provided to Versus (also installed at ECMWF) for the comparison of the 2 model versions with the computation of scores and plots <u>at both resolutions</u>.

Activity during the COSMO year

- **September-October-November 2015**
  - Test of COSMO v5.03 and comparison against v5.01 at 7 km.
  - 2.5 months were taken to run the experiments, perform verification on Versus, produce the report: TOO LONG!
  - Some critical issues were raised and room for improvement in some areas was identified.
- **May-June 2016**
  - Test of COSMO v5.04a and comparison against v5.03 **at both 7 km and 2.8 km**
  - "Misunderstanding" on the setting of the namelists to be used at 2.8 km ... The runs were almost useless (some "hot" emails at the end of June), BUT
  - we were much faster than before!!!

NWP METEOROLOGICAL TEST SUITE:

integration domain (for both 7 and 2.8 km)

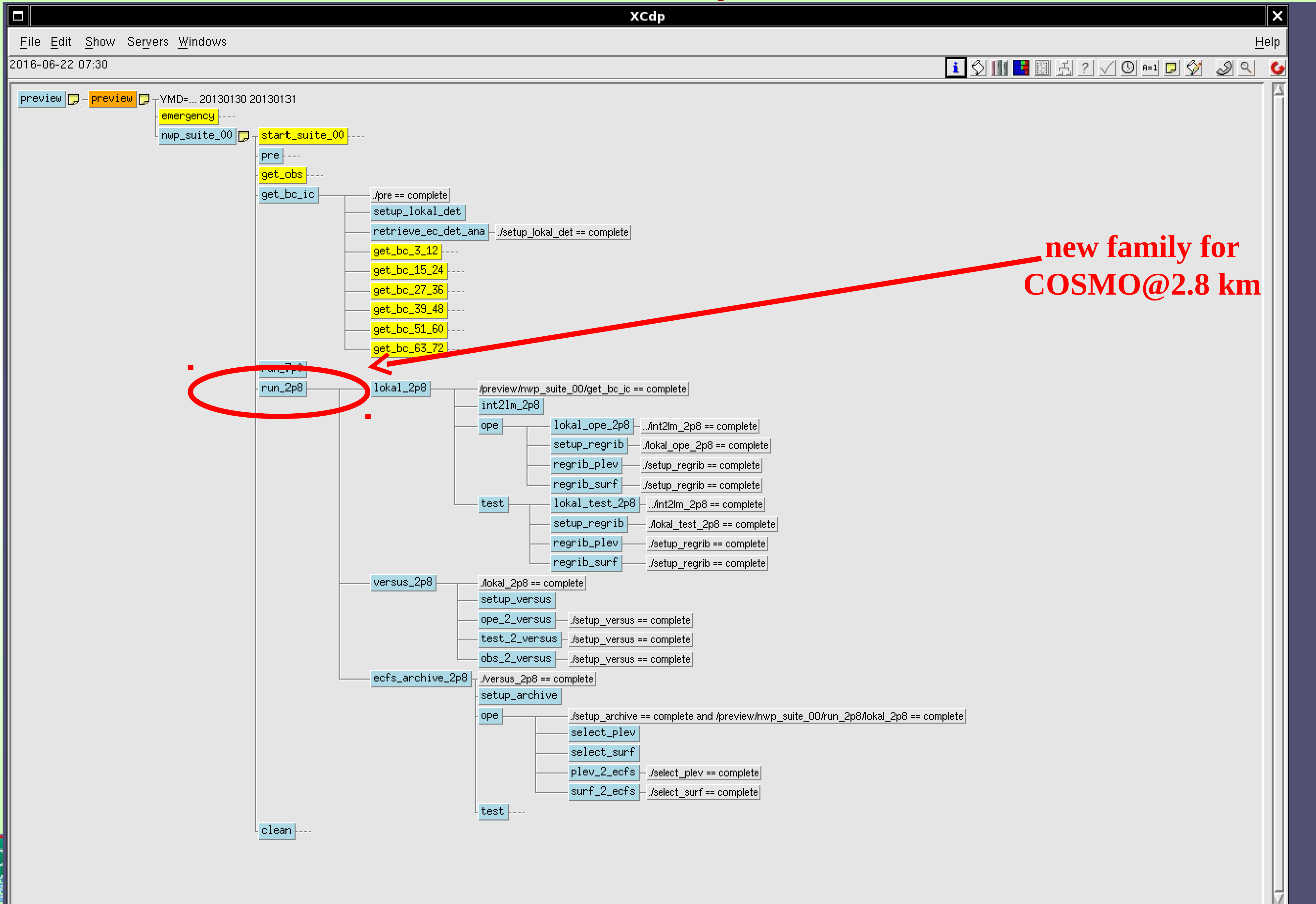


ECMWF HRES: ec_nx = 801; ec_ny = 401; 137 ML; ec_dlon = ec_dlat = 0.125 (**14 km**); fc+72h

COSMO@7p0: ie_tot = 745 ; je_tot = 569; 40 ML; dlon = dlat = 0.625 **(7 km)**; fc+72h

COSMO@2p8: ie_tot = 1799 ; je_tot = 1369; 50 ML; dlon = dlat = 0.025 **(2.8 km)**; fc+48h

Screenshot of the suite with its main families/tasks



Performances and costs at *7p0* and 2p8

int2lm

Interpolation for COSMO-5.04a (HRES --> COSMO@7.0) 330 sec , 43 SBU
total_tasks and node for int2lm (@7p0): EC_total_tasks=36, EC_nodes=1

Interpolation for COSMO-5.03 (HRES --> COSMO@2.8) 864 sec, 278 SBU
Interpolation for COSMO-5.04a (HRES --> COSMO@2.8) 864 sec, 278 SBU
total_tasks and nodes for int2lm (@2p8): EC_total_tasks=72, EC_nodes=2

COSMO

COSMO-5.03 @7.0 928 sec, 2993 SBU
COSMO-5.04a @7.0 " "
total_tasks and nodes for COSMO (@7p0): EC_total_tasks=720, EC_nodes=20

COSMO-5.03 @2.8 6616 sec, 38417 SBU
COSMO-5.04a @2.8 6145 sec, 35682 SBU (COSMO@2.8 km is very expensive!)
total_tasks and nodes for COSMO (@2p8): EC_total_tasks=1296, EC_nodes=36

Open issues

Room for improvement:

- The full chain of the COSMO NWP suite is rarely run (mostly, twice a year); every time, you need to remember what you did last time. The suite requires the involvement of COSMO scientists working in 6 different institutions (DWD, Arpae-SIMC, ARPA-Piedmont, COMET, NMA, HNMS).
  - good side: example of collaboration and synergy of expertise within COSMO;
  - bad side: coordination of work is extremely difficult and timeliness remains a dream...

Shortage of Billing Units in ECMWF Special Project (SPITRASP):

- In May-June 2016, ECMWF upgraded the processors of the super-computers. COSMO is about 1.5 more expensive on the new processors (we could not know this last year!). We have already spent 4.8 million BUs out of the 5.0 millions allocation for 2016.
  - On 24/8, we applied for extra-resources to test next model release.

COSMO-LEPS upgrade

COSMO-LEPS suite @ ECMWF: present status

d-1 d d+1 d+2 d+3 d+4 d+5

older EPS

00

younger EPS

12

clustering period

4 variables

Z U V Q

3 levels

500 700 850 hPa

Cluster Analysis and RM identification

2 time steps

European area

Complete Linkage

16 Representative Members driving the 16 COSMO integrations (weighted according to the cluster populations)

using Tiedtke convection scheme

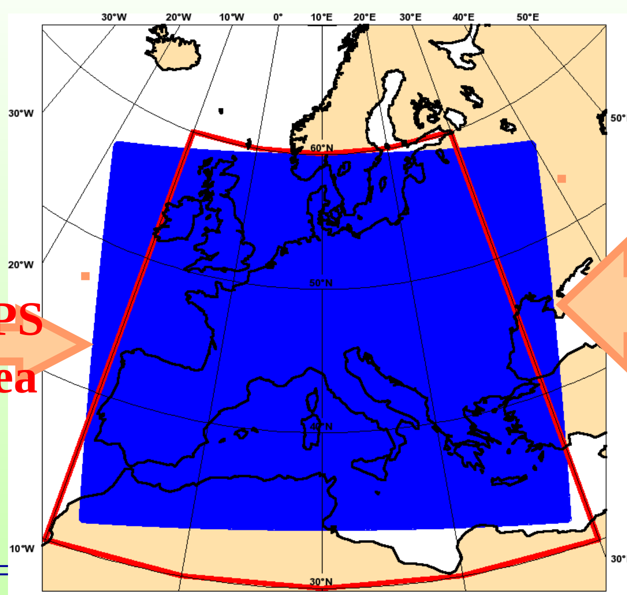
+

perturbations in turbulence scheme and in physical parameterisations

+

soil initial conditions from COSMO-EU

COSMO-LEPS clustering area



COSMO-LEPS Integration Domain

- suite runs as a "time-critical application" managed by ARPA-SIMC; runs at both 00 and 12TC;
- Δx ~ 7 km; 40 ML; fc+132h;
- COSMO v5.03 since 1 February 2016,
- computer time (57 million BUs for 2016) provided by the COSMO partners which are ECMWF member states.

Types of perturbations (2016)

As for types and values, the results from CSPERT experimentation were followed (* denotes default values for COSMO):

- convection_scheme: Tiedtke* (members 1-16), ←
- tur_len (either 150, or 500*, or 1000),
- pat_len (either 500*, or 2000),
- crsmin (either 50, or 150*, or 200),
- rat_sea (either 1, or 20*, or 40),
- rlam_heat (either 0.1, or 1*, or 5),
- mu_rain : either 0.5* (with rain_n0_factor =0.1) or 0 (with rain_n0_factor =1.0),
- cloud_num (either 5x10^8* or 5x10^7).

Recent news_1

- **December 2015**
  - 30-day tests of COSMO-LEPS with ICON-EU soil fields: no noticeable impact on short-range forecast skill of TP, T2M, TD2M.
- **1 February 2016: suite upgrade**
  - COSMO version update (5.01 → 5.03); int2lm 2.0;
  - Production and archive of 100 metre U and V wind component;
  - Archive of P, T, U, V at model levels 35, 36, 37, 38, 39, 40.
- **19 February 2016: int2lm**
  - ECMWF fields (from test dissemination) with longitudeOfFirstGridPoint = 335000 (instead of longitudeOfFirstGridPoint = -25000) made int2lm fail;
  - a patch was applied to handle ECMWF GRIB1 files with longitudes greater than 180°.
- **25 February 2016: field production to ARPA-Liguria**
  - Dissemination of COSMO-LEPS fields in GRIB2 format.

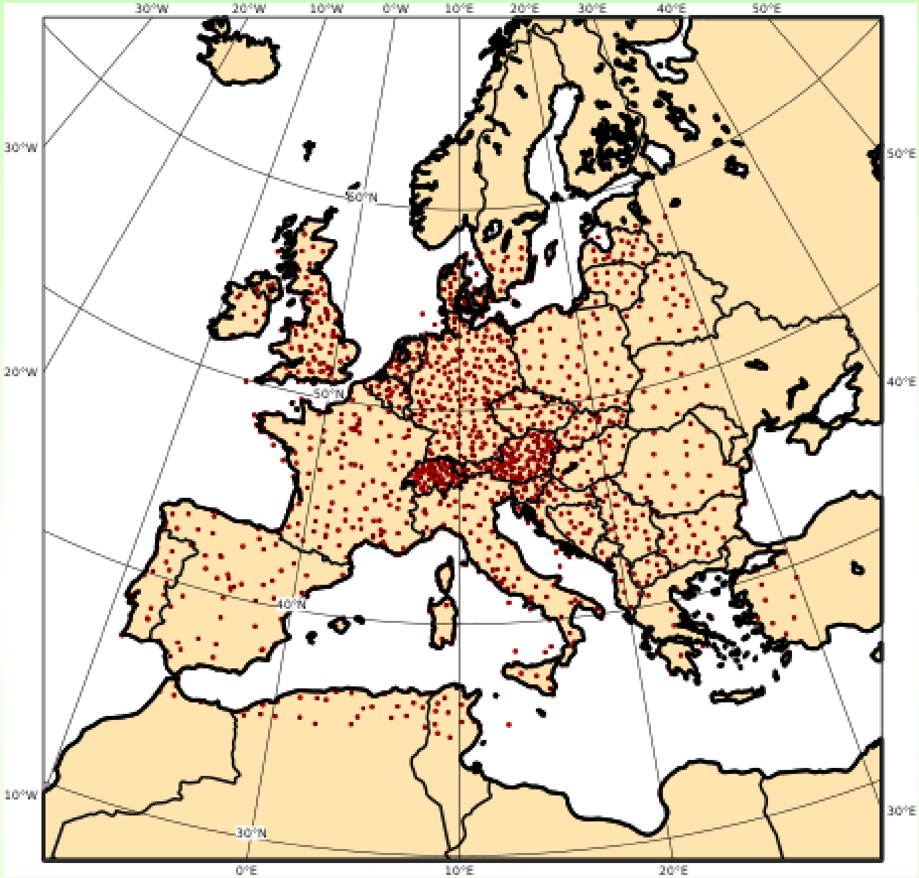
Recent news_2

- **6 June 2016: ECMWF upgrade**
  - Change of processors on ECMWF super-computers (from IvyBridge to Broadwell) ➔ change of geometry in COSMO and int2lm configurations; no impact on users; change of costs.
- **11 June 2016: beginning of esuite**
  - Start of experimentation of COSMO-LEPS with 20 members in single precision (**20_sp**) and comparison against operational COSMO-LEPS (16 members in double precision, **16_dp**).
    - Meteorological aspects
    - Computational aspects

Meteorological aspects

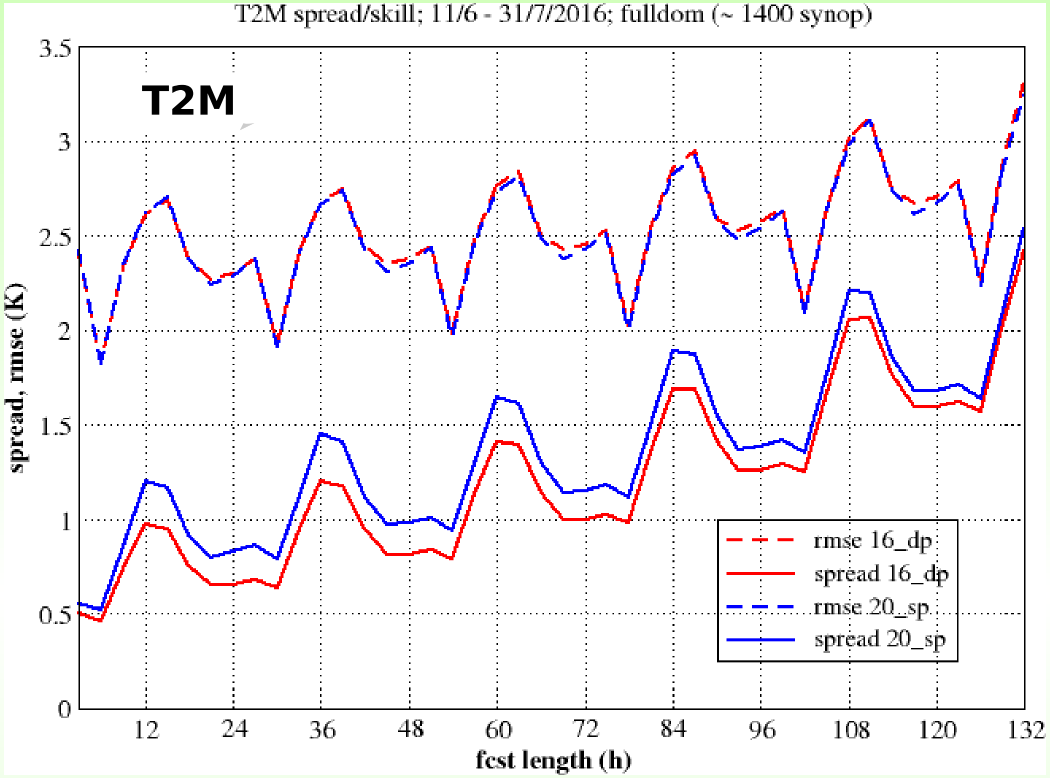
- COSMO v5.03: inter-comparison of **16_dp (no SPPT)** and **20_sp (with SPPT) .**
- Same soil initial conditions from COSMO-EU.
- Both the cluster analyses and the random choice of perturbation parameters are performed separately for 16_dp and 20_sp.
- **51 days of test (from 11/6 to 31/7/2016)**, starting at 00UTC.
- Consider performance in terms of:
    - 2-metre temperature,
    - 10-metre wind-speed,
    - 12-hour cumulated precipitation.

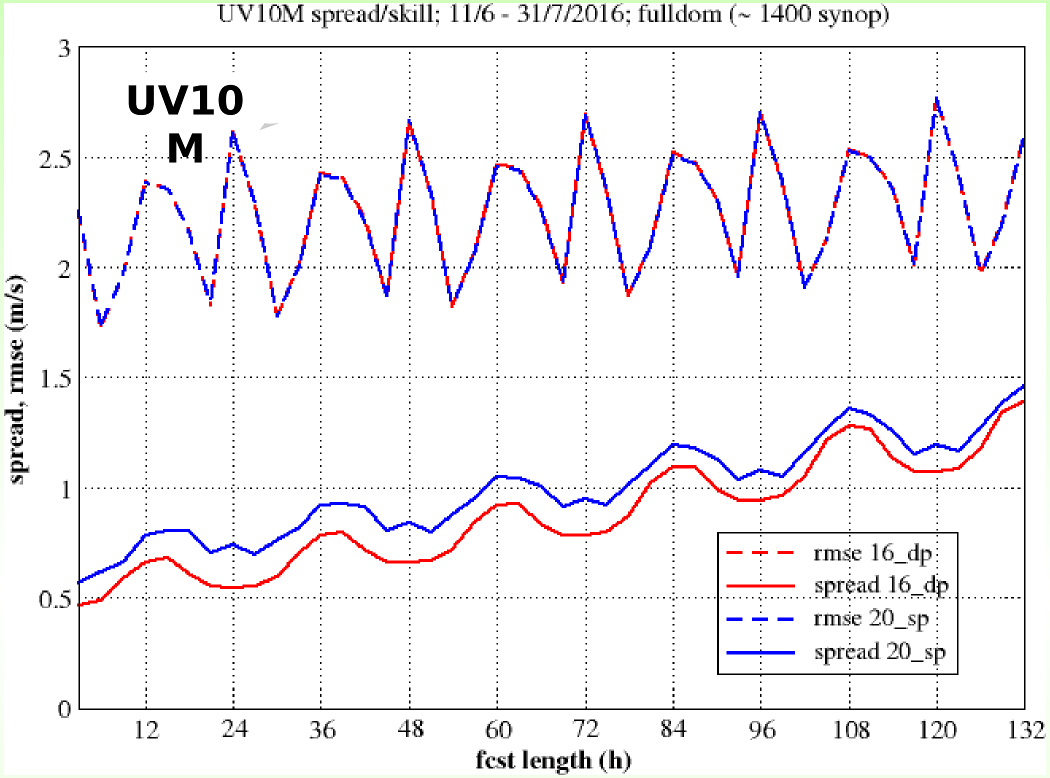
    (thresholds:1, 5, 10, 15, 25, 50 mm/12h).



Verification area: full domain (~ 1400 synop reports).

Spread/skill for T2M and UV10M



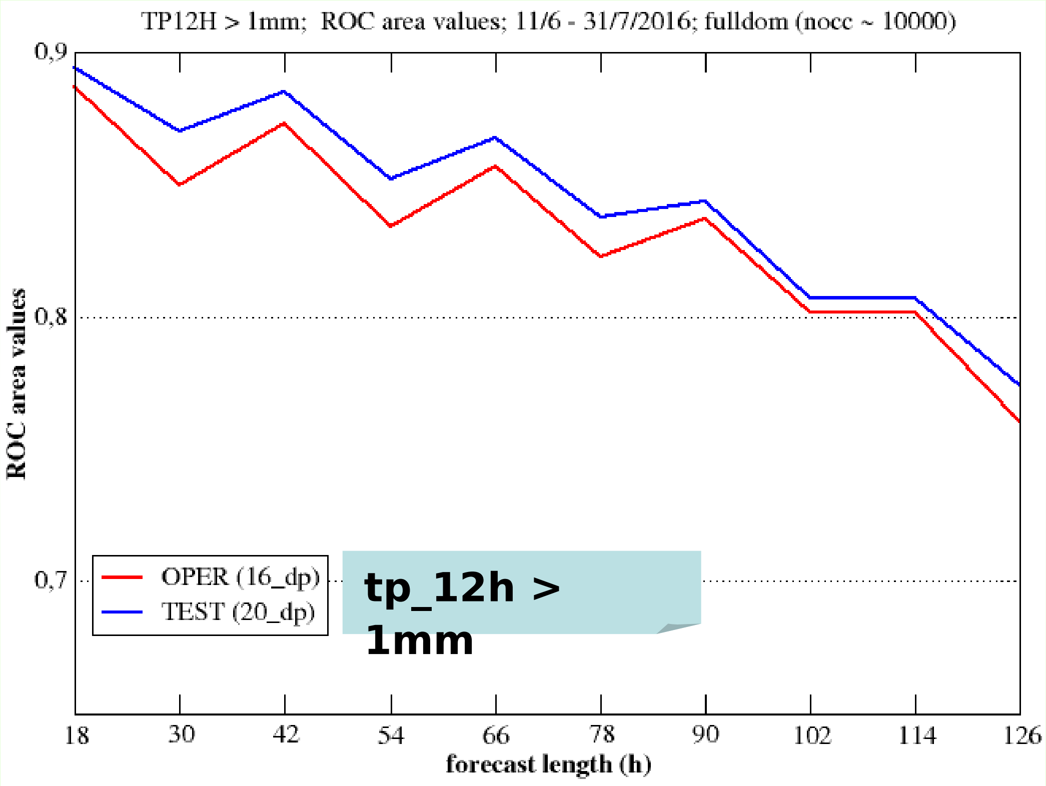


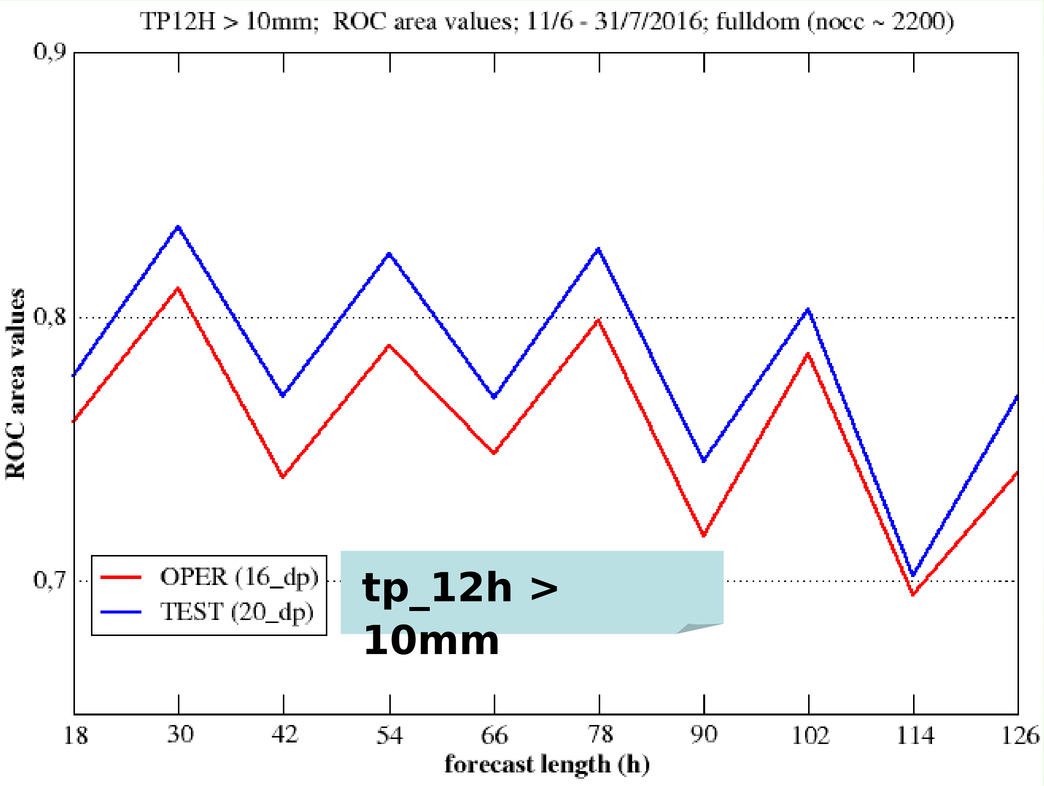
- **Larger spread for 20_sp for both variables**; in either cases, lack of spread in the short range.
- T2M: the daily cycle of the spread follows to a certain extent the cycle of the error.
- **Limited impact** (if any) **on the forecast skill** of the ensemble mean.

It seems we are going in the right direction.

Probabilistic prediction of tp: ROC area

- Area under the curve in the HIT rate vs FAR diagram; the higher, the better …
- Valuable forecast systems have ROC area values > 0.6.
- Consider two events: 12-hour precipitation exceeding 1 mm and 10 mm.

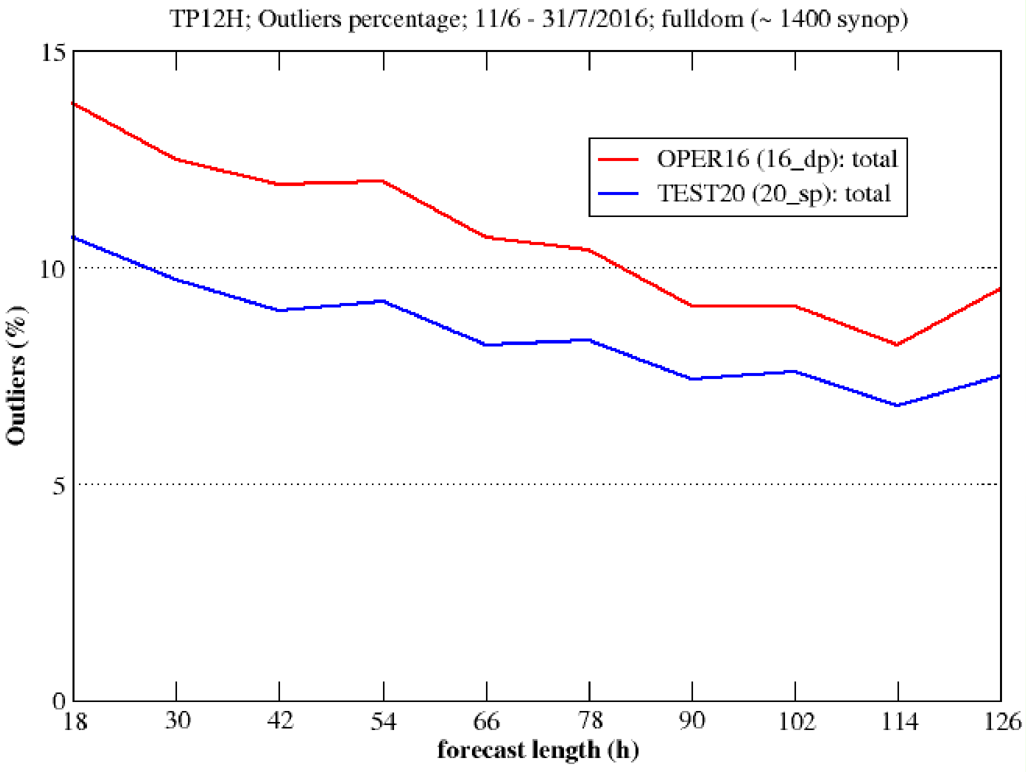


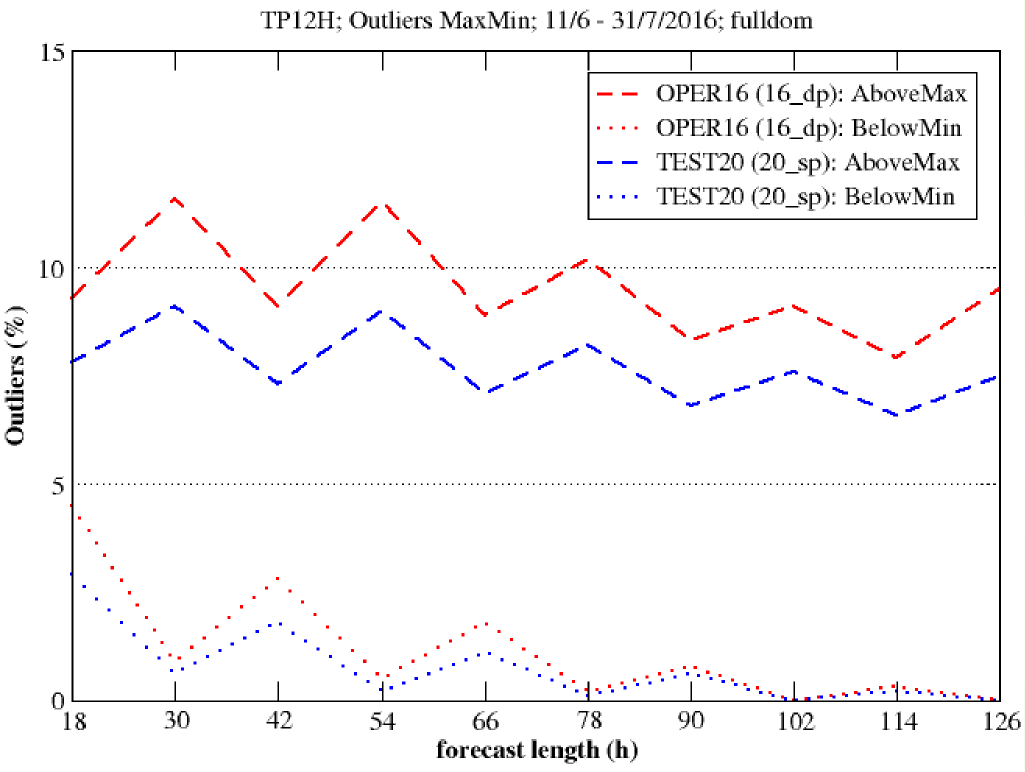


- Better performance by **20_sp** for both thresholds.
- Impact more evident in the short range.

Probabilistic prediction of tp: OUTLIERS

- How many times the analysis is out of the forecast interval spanned by the ensemble members.
- … the lower the better …





- Reduction of outliers for **20_sp** COSMO-LEPS for all forecast ranges.
- Decrease of outliers especially AboveMax (in 20_sp, it happens less frequently that all ensemble members predict lower precipitation than what observed).

Computational aspects

#PBS –l EC_nodes=20

#PBS –l EC_total_tasks=720

………

what do you gain if you run in single precision?

Last year (with COSMO v5.1 and old ECMWF processors with different geometry):

the gain was highly variable from day to day (min: ~10%; max: ~50%), but

➔ *average saving: ~ 35% x run*

this year (with COSMO v5.03)

	Elapsed time (sec)	Cost of 1 member (ECMWF BU)
double_precision	**960**	**3100**
single_precision	**500**	**1600**

➔ **average saving: ~ 48%** x run (impact of SPPT is negligible in terms of computer time)

→ **16_dp cost ~ 49600 BU**

→ **20_sp cost ~ 32000 BU**

Despite the 25% increase in ensemble size, **20_sp** is still **35% cheaper** than **16_dp**!!!

But…

Frequent explosions of COSMO in single precision with SPPT (5-6 crashes every day!)

- Namelist changes did not cure the problem with COSMO v5.03.
- Need of code modifications (bug fix in divergence damping, targeted diffusion to prevent significant temperature anomalies → l_diff_cold_pools=.true.) not yet available in the official release.

“Plaster” during experimentations:

When the task failed, COSMO was resubmitted with SPPT=.false. (and then the task ran successfully).

COSMO-LEPS with SPPT: namelist

```
&RUNCTL
  ....
  leps    =.TRUE.,
  lsppt   =.TRUE.,
/END
```

```
cat >! $workingDir/INPUT_EPS << EONL
 &EPSCTL
  iepsmem=$MEMBER,
  iepstot=$LM_NL_EPSMEMBERS,
  iepstyp=203
  imode_rn=1,
  itype_vtaper_rn=2,
  itype_qxpert_rn=2,
  itype_qxlim_rn=0,
  npattern_rn=1,
  hinc_rn=6,
  dlat_rn=5.0,
  dlon_rn=5.0,
  stdv_rn=1.0,
  range_rn=0.9,
  lgauss_rn=.TRUE.,
  lhorint_rn=.TRUE.,
  ltimeint_rn=.TRUE.,
 /END
EONL
```

Results and road-map

20_sp COSMO-LEPS:

- has better spread/skill relation for temperature and wind-speed,
- provides more accurate probabilistic prediction of precipitation,
- is cheaper,
- is faster

than the operational system (**16_dp COSMO-LEPS**).

SON2016: perform a few more experiments and start test-dissemination of 20_sp COSMO-LEPS.

- For the moment, use SPPT=.false.; once COSMO v5.05 is available, the explosion problems should be fixed and we can switch to SPPT=.true.
- **Go operational before Christmas 2016?!?!!?**

Future work

- By the end of October: migrate from COSMO-EU to Icon-Regional for the provision of soil-moisture analysis fields.
- Use high resolution boundaries from ECMWF ENS (already tested).
- Implement and use weighted products (e.g. weighted ensemble mean).
- Upgrade Fieldextra.
- Listen to users.
- ....

THANKS FOR YOUR ATTENTION

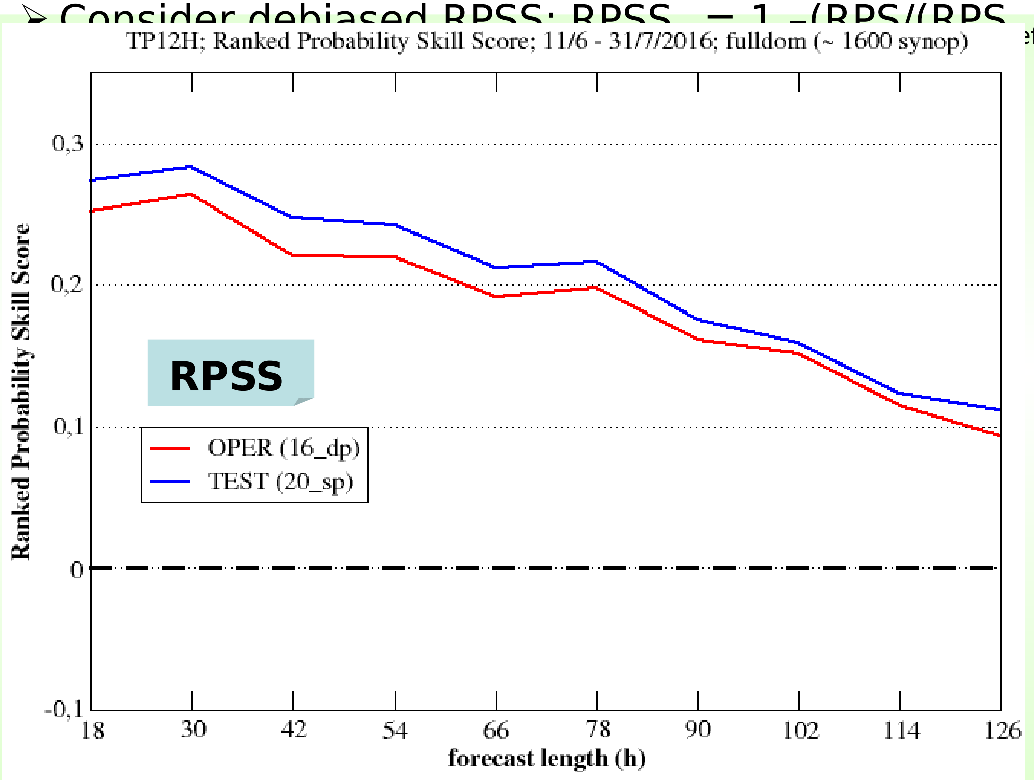
But...

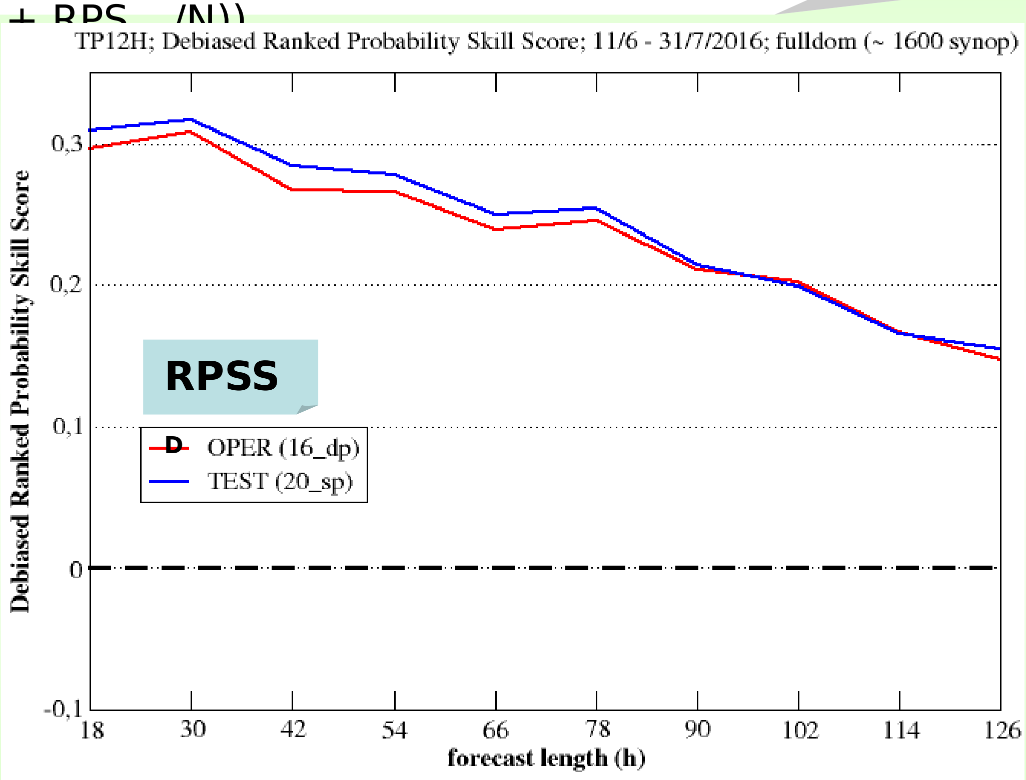
Frequent explosions (5-6 every day!) of COSMO in single precision with SPPT.

- Namelist changes did not cure the problem with COSMO v5.03.
- Plaster for experiments: when the task failed, COSMO was resubmitted with SPPT=.false. (and the task ran successfully).
- Need of code modifications (bug fix in divergence damping, targeted diffusion to prevent significant temperature anomalies l_diff_cold_pools=.true.) not yet available in the official release.

Probabilistic prediction of tp: RPSS

- BSS "cumulated" over all thresholds. RPSS is written as 1-RPS/RPS$_{ref}$. **Sample climate** is the reference system. RPS is the extension of the Brier Score to the multi-event situation.
- Useful forecast systems for RPSS > 0; RPSS depends on the ensemble size, penalising small ensemble sizes.
- Consider debiased RPSS: RPSS$_D$ = 1 –(RPS/(RPS$_{ref}$ + RPS$_{ref}$/N))

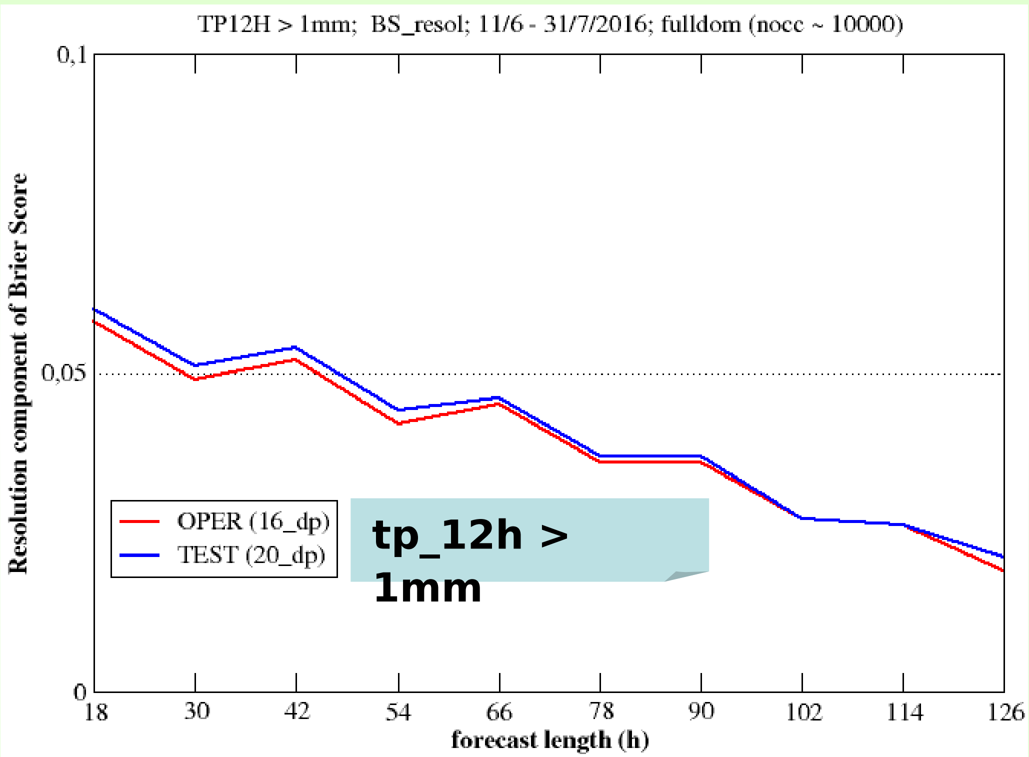


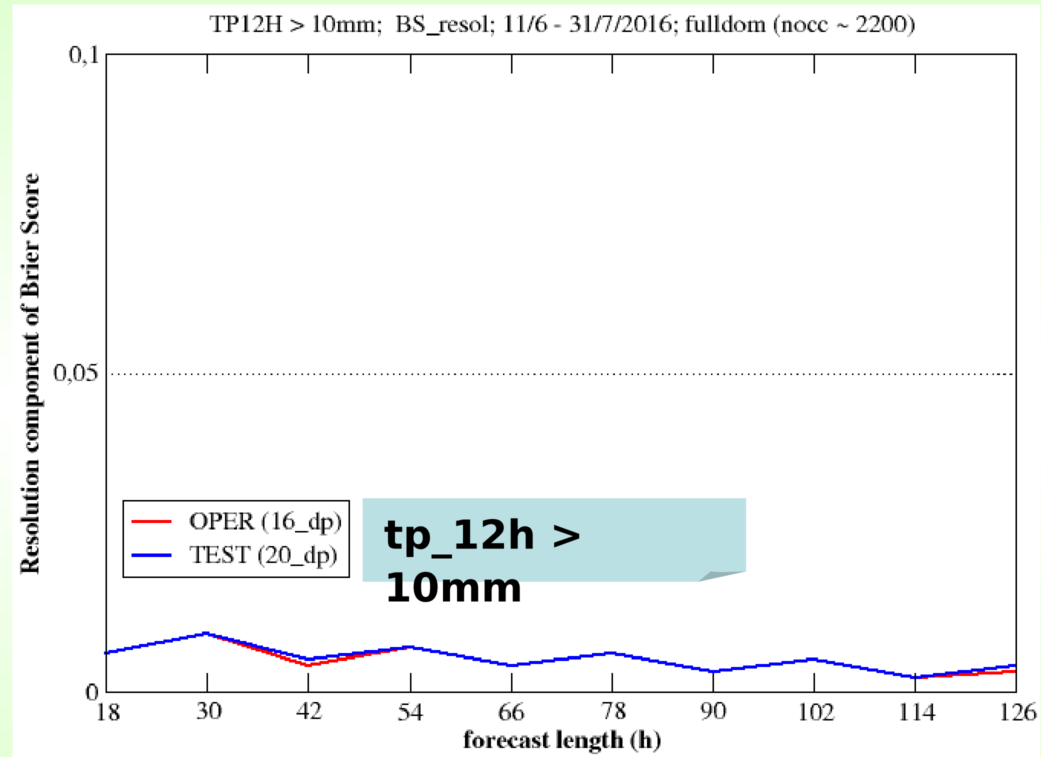


• In either cases (RPSS or RPSS$_D$), better performance of **20_sp** COSMO-LEPS, more evident for short ranges.

Probabilistic prediction of tp: Resolution

- Resolution component of the Brier Score: describes the ability of the system to distinguish among events in different categories; the higher, the better …
- Consider two events: 12-hour precipitation exceeding 1 mm and 10 mm.



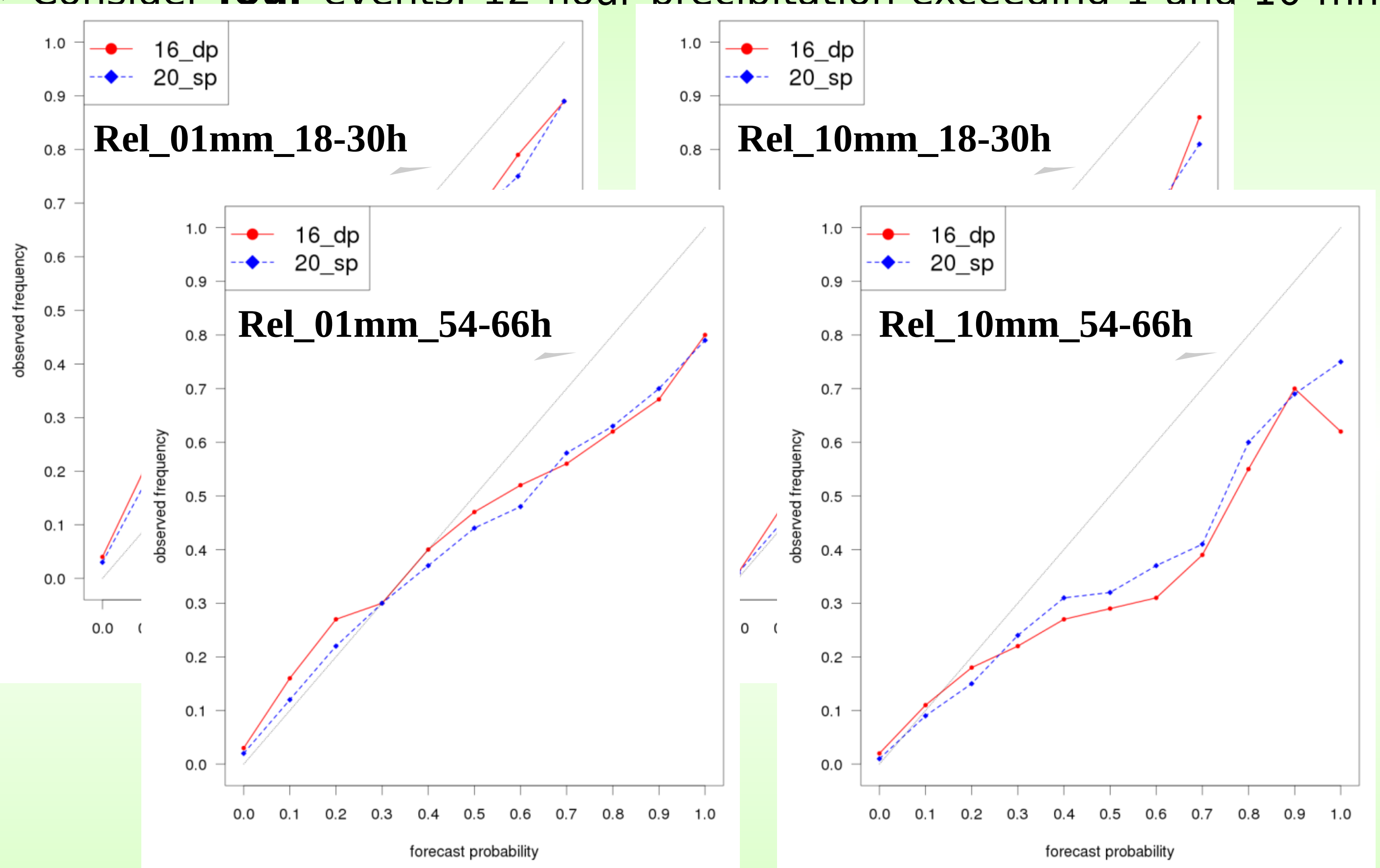


- Slightly better performance by **20_sp** only for the lower threshold.
- Impact more evident in the short range for 1mm threshold.

Probabilistic prediction of tp: Reliability

➢Match between fcst probability and obs frequency for a certain event; the closer to the diagonal, the better ....

➢Consider **four** events: 12-hour precipitation exceeding 1 and 10 mm at the ranges 18-30h and 54-



• COSMO-LEPS overconfidence increases with both threshold and forecast range (fcst_prob > obs_freq) for both **16_dp** and **20_sp**.

• Not clear positive impact of enlarged ensemble size.

Computational aspects

#PBS –l EC_nodes=20

#PBS –l EC_total_tasks=720

………

what do you gain if you run in single precision?

Last year (with COSMO v5.1 and old ECMWF processors with different geometry) :

the gain was highly variable from day to day (min: ~10%; max: ~50%), but

➔*average saving: about 35% x run*

THIS YEAR (COSMO v5.03)

	double precision	**single precision**
• Cost of 1 COSMO-LEPS run (ECMWF Billing Units)	3100	1600
• Elapsed time (sec)	960	500

➔ **average saving: about 48%** x run

(impact of SPPT is negligible in terms of computer time)

Despite the 25% increase in ensemble size, **20_sp is cheaper than 16_dp!!!**

Tested configurations

- **ECMWF IFS:**

ec_nx = 801; ec_ny = 401; ec_dlon = ec_dlat = 0.125 (15 km); 137 ML; fc+72h

- **COSMO_7**

ie_tot = 745 ; je_tot = 569 ; dlon = dlat = 0.625 **(7 km);** 40 ML; fc+72h

- **COSMO_28**

ie_tot = 1799 ; je_tot = 1369 ; dlon = dlat = 0.025 **(2.8 km);** 50 ML; fc+48h

PT NWP METEOROLOGICAL TEST SUITE

Task 1: COSMO Model Installation and Implementation

1.2 Model Implementation and Set-up of Appropriate Tests

TESTS:

- 72 hours run
- one daily cycle based on 00UTC initializing data
- ECMWF inital and LBC (11.6Gb)
- Runs for January 2013 and July 2013 **(62 days in total)**

COSTS (on IBM):

- Interpolation for **COSMO-4.26**: **~ 80.0 BU** per run (takes ~ 8 min)
- Interpolation for **COSMO-5.0**: **~ 81.5 BU** per run (takes ~ 8 min)
- **COSMO-4.26**: **~ 2434 BU** per run (takes ~ 30 min)
- **COSMO-5.0**: **~ 2350 BU** per run (takes ~ 29 min)
- **COSMO-5.1: ~ 2284 BU per run (takes ~ 28 min) !!!!!!!!!!!!!!!!!**
- total_tasks = 64 and node = 1 for **int2lm**
- total_tasks = 512 and node = 8 for **COSMO**

Task 1: COSMO Model Installation and Implementation

A. Montani, R.C. Dumitrache

1.1 Development of the Test Suite

ECMWF resources – special project SPITRASP (submitted by A. Raspanti)

Computer resources	2013		2014		2015
	Allocated	Used	Allocated	Used	Allocated
HPC Facility (units)	400 000	11.91	1 000 000	356 420.40	1 000 000
Data storage capacity (GB)	80	1	180	20	180

Most (all?) of COSMO development is devoted towards the improvement of the model at the convection-permitting scale (below 3 km of horizontal resolution).

==> It makes sense to compare the skill of the new model versions at such resolution.

Need to consider the extension to 2.8 km

There are (at least) two options for the implementation of the high-resolution part:

Option A: **IFS → COSMO_7 → COSMO_28**

Option B: **IFS → COSMO_28**

ECMWF IFS: ec_nx=801; ec_ny=401; 137 ML; ec_dlon=ec_dlat = 0.125 (**14 km**); fc+72h

COSMO_7: ie_tot = 745 ; je_tot = 569; 40 ML; dlon = dlat = 0.625 **(7 km)**; fc+72h

COSMO_28: ie_tot = 1799 ; je_tot = 1369; 50 ML; dlon = dlat = 0.025 **(2.8 km)**; fc+48h

Final remarks

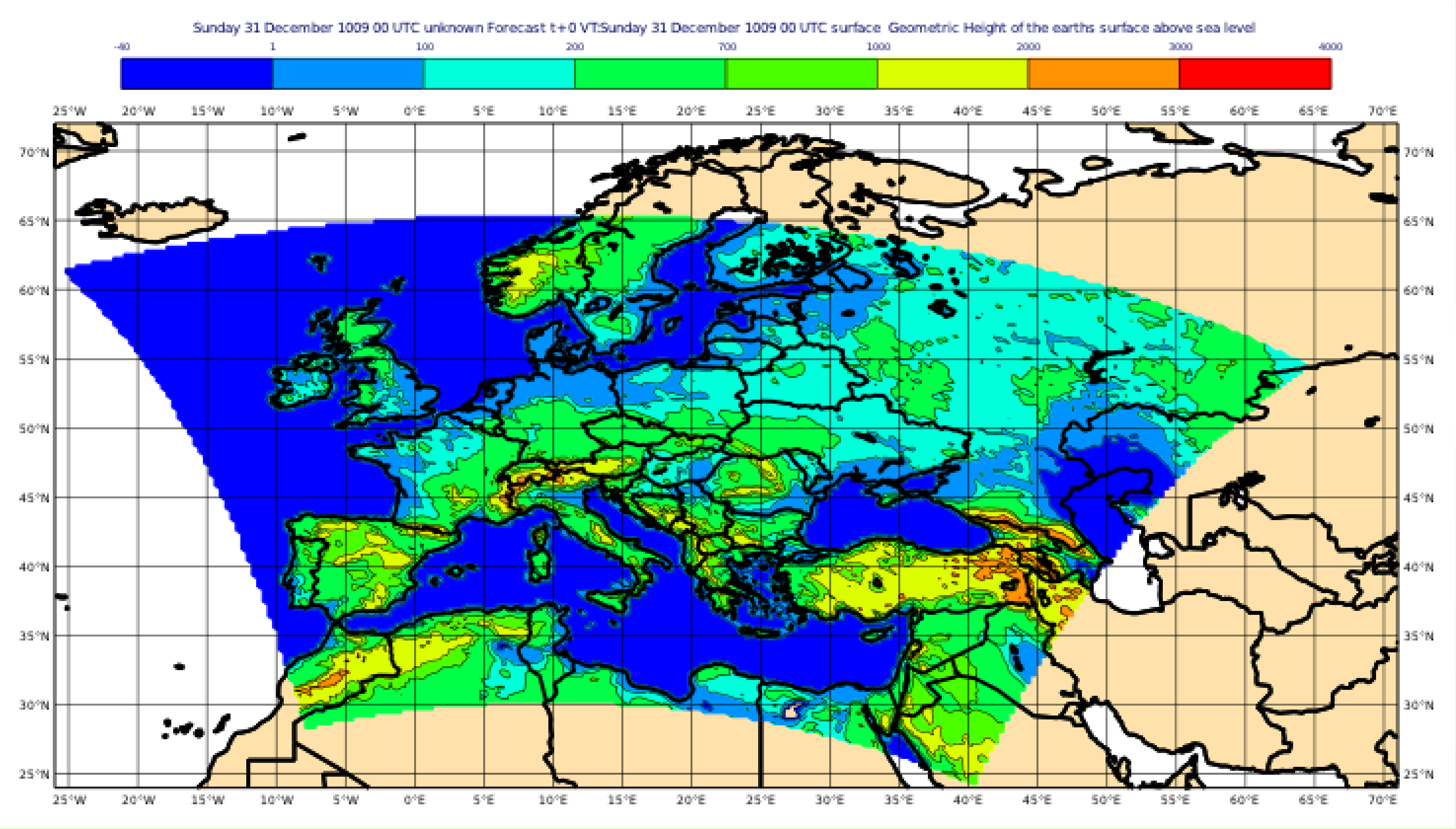
Option A: **IFS → COSMO_7 → COSMO_28**

Option B: **IFS → COSMO_28**

move, anyway, to high resolution

- Option A enables a very comprehensive verification.
- Option A has probably more impact on the “verification” people, as both COSMO_7 and COSMO_28 should be loaded and verified by Versus @ ECMWF.
- Option B is simpler to implement/ maintain.
- Option B is cheaper in terms of needs of computing resources.
- Option B might require a higher-resolution verification network (that we don’t have!).

Integration domain at 2.8km (almost the same as 7km)



CONSORTIUM FOR SMALL SCALE MODELING